

## Claims

- Sub B)*
- [c1] 1. A method for automatically generating a query, comprising:  
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;  
identifying a set of entities in selected document content for searching additional information related thereto using the information retrieval system;  
assigning the selected document content a classification label from the organized classification of content;  
automatically formulating a query that restricts a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.
- [c2] 2. The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.
- [c3] 3. The method according to claim 2, wherein the number of terms added is limited to a predefined number.
- [c4] 4. The method according to claim 2, further comprising limiting the query by adding terms defining the assigned classification label.
- [c5] 5. The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.
- [c6] 6. The method according to claim 1, further comprising using a text categorizer to assign the classification label assigned from the organized classification of content.
- [c7] 7. The method according to claim 6, further comprising:  
extracting with the text categorizer a set of terms relating to the document content; and  
appending to the query ones of the set of terms extracted by the text categorizer to contextualize the query.
- [c8] 8. The method according to claim 7, further comprising abbreviating the set of terms extracted by the text categorizer to a predefined limit of terms.

- [c9] 9. The method according to claim 8, wherein said abbreviating comprises:  
extracting noun phrases from the selected document content;  
ranking the noun phrases by those that occur most frequently in the document content;  
defining a subset of noun phrases by identifying those ranked noun phrases that occur more frequently than a first predefined frequency;  
ranking those words in the subset of noun phrases by their frequency of occurrence to define an ordered list of words;  
defining a subset of the ordered list of words by identifying those ranked words that occur more frequently than a second predefined frequency;  
re-ranking the subset of words in inverse frequency to their use in the category of information in the information retrieval system identified by the assigned classification label;  
using only those highest ranked words in the re-ranked subset of words to define the set of terms appended to the query.
- [c10] 10. The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.
- [c11] 11. The method according to claim 10, further comprising ranking results from the query performed at the information retrieval system in accordance with one of the assigned classification label and the characteristic vocabulary.
- [c12] 12. The method according to claim 11, using the method in a system for enriching selected content of a document with personalities that identify enrichment themes.
- [c13] 13. The method according to claim 1, further comprising automatically identifying the set of entities using a service that recognizes entities of a predefined type.
- [c14] 14. A system for automatically generating a query, comprising:  
an entity extractor for identifying a set of entities in selected document content for searching information related thereto using an information retrieval system;  
a categorizer for defining an organized classification of document content with each class in the organization of content having associated therewith a classification label; each classification label corresponding to a category of information in the information retrieval system; the categorizer assigning the selected document content a classification label from the organized classification of content;

a query generator for automatically formulating a query that restricts a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.

[c15] 15. The system according to claim 14, further comprising a short length aspect vector generator for generating terms relating to context information surrounding the set of entities in the selected document content; wherein the query generator adds the terms relating to the context information to limit the query.

[c16] 16. The system according to claim 15, wherein the query generator further limits the query by adding terms defining the selected classification label provided by the categorizer.

[c17] 17. The system according to claim 16, further comprising a content manager for enriching the selected document content with results provided from the information retrieval system using the query.

[c18] 18. An article of manufacture for use in a computer system, comprising:  
a memory;  
instructions stored in the memory for operating a method for automatically generating a query, comprising:  
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;  
identifying a set of entities in selected document content for searching information related thereto using the information retrieval system;  
assigning the selected document content a classification label from the organized classification of content;  
automatically formulating a query that restricts a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.

[c19] 19. The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context

[c20]

information surrounding the set of entities in the selected document content.

20. The article of manufacture according to claim 19, wherein the instructions stored in the memory further comprise further limiting the query by adding terms defining the assigned classification label.

6  
10  
14  
18  
22  
26  
30  
34  
38  
42  
46  
50  
54  
58  
62  
66  
70  
74  
78  
82  
86  
90  
94  
98  
102  
106  
110  
114  
118  
122  
126  
130  
134  
138  
142  
146  
150  
154  
158  
162  
166  
170  
174  
178  
182  
186  
190  
194  
198  
202  
206  
210  
214  
218  
222  
226  
230  
234  
238  
242  
246  
250  
254  
258  
262  
266  
270  
274  
278  
282  
286  
290  
294  
298  
302  
306  
310  
314  
318  
322  
326  
330  
334  
338  
342  
346  
350  
354  
358  
362  
366  
370  
374  
378  
382  
386  
390  
394  
398  
402  
406  
410  
414  
418  
422  
426  
430  
434  
438  
442  
446  
450  
454  
458  
462  
466  
470  
474  
478  
482  
486  
490  
494  
498  
502  
506  
510  
514  
518  
522  
526  
530  
534  
538  
542  
546  
550  
554  
558  
562  
566  
570  
574  
578  
582  
586  
590  
594  
598  
602  
606  
610  
614  
618  
622  
626  
630  
634  
638  
642  
646  
650  
654  
658  
662  
666  
670  
674  
678  
682  
686  
690  
694  
698  
702  
706  
710  
714  
718  
722  
726  
730  
734  
738  
742  
746  
750  
754  
758  
762  
766  
770  
774  
778  
782  
786  
790  
794  
798  
802  
806  
810  
814  
818  
822  
826  
830  
834  
838  
842  
846  
850  
854  
858  
862  
866  
870  
874  
878  
882  
886  
890  
894  
898  
902  
906  
910  
914  
918  
922  
926  
930  
934  
938  
942  
946  
950  
954  
958  
962  
966  
970  
974  
978  
982  
986  
990  
994  
998  
1002  
1006  
1010  
1014  
1018  
1022  
1026  
1030  
1034  
1038  
1042  
1046  
1050  
1054  
1058  
1062  
1066  
1070  
1074  
1078  
1082  
1086  
1090  
1094  
1098  
1102  
1106  
1110  
1114  
1118  
1122  
1126  
1130  
1134  
1138  
1142  
1146  
1150  
1154  
1158  
1162  
1166  
1170  
1174  
1178  
1182  
1186  
1190  
1194  
1198  
1202  
1206  
1210  
1214  
1218  
1222  
1226  
1230  
1234  
1238  
1242  
1246  
1250  
1254  
1258  
1262  
1266  
1270  
1274  
1278  
1282  
1286  
1290  
1294  
1298  
1302  
1306  
1310  
1314  
1318  
1322  
1326  
1330  
1334  
1338  
1342  
1346  
1350  
1354  
1358  
1362  
1366  
1370  
1374  
1378  
1382  
1386  
1390  
1394  
1398  
1402  
1406  
1410  
1414  
1418  
1422  
1426  
1430  
1434  
1438  
1442  
1446  
1450  
1454  
1458  
1462  
1466  
1470  
1474  
1478  
1482  
1486  
1490  
1494  
1498  
1502  
1506  
1510  
1514  
1518  
1522  
1526  
1530  
1534  
1538  
1542  
1546  
1550  
1554  
1558  
1562  
1566  
1570  
1574  
1578  
1582  
1586  
1590  
1594  
1598  
1602  
1606  
1610  
1614  
1618  
1622  
1626  
1630  
1634  
1638  
1642  
1646  
1650  
1654  
1658  
1662  
1666  
1670  
1674  
1678  
1682  
1686  
1690  
1694  
1698  
1702  
1706  
1710  
1714  
1718  
1722  
1726  
1730  
1734  
1738  
1742  
1746  
1750  
1754  
1758  
1762  
1766  
1770  
1774  
1778  
1782  
1786  
1790  
1794  
1798  
1802  
1806  
1810  
1814  
1818  
1822  
1826  
1830  
1834  
1838  
1842  
1846  
1850  
1854  
1858  
1862  
1866  
1870  
1874  
1878  
1882  
1886  
1890  
1894  
1898  
1902  
1906  
1910  
1914  
1918  
1922  
1926  
1930  
1934  
1938  
1942  
1946  
1950  
1954  
1958  
1962  
1966  
1970  
1974  
1978  
1982  
1986  
1990  
1994  
1998  
2002  
2006  
2010  
2014  
2018  
2022  
2026  
2030  
2034  
2038  
2042  
2046  
2050  
2054  
2058  
2062  
2066  
2070  
2074  
2078  
2082  
2086  
2090  
2094  
2098  
2102  
2106  
2110  
2114  
2118  
2122  
2126  
2130  
2134  
2138  
2142  
2146  
2150  
2154  
2158  
2162  
2166  
2170  
2174  
2178  
2182  
2186  
2190  
2194  
2198  
2202  
2206  
2210  
2214  
2218  
2222  
2226  
2230  
2234  
2238  
2242  
2246  
2250  
2254  
2258  
2262  
2266  
2270  
2274  
2278  
2282  
2286  
2290  
2294  
2298  
2302  
2306  
2310  
2314  
2318  
2322  
2326  
2330  
2334  
2338  
2342  
2346  
2350  
2354  
2358  
2362  
2366  
2370  
2374  
2378  
2382  
2386  
2390  
2394  
2398  
2402  
2406  
2410  
2414  
2418  
2422  
2426  
2430  
2434  
2438  
2442  
2446  
2450  
2454  
2458  
2462  
2466  
2470  
2474  
2478  
2482  
2486  
2490  
2494  
2498  
2502  
2506  
2510  
2514  
2518  
2522  
2526  
2530  
2534  
2538  
2542  
2546  
2550  
2554  
2558  
2562  
2566  
2570  
2574  
2578  
2582  
2586  
2590  
2594  
2598  
2602  
2606  
2610  
2614  
2618  
2622  
2626  
2630  
2634  
2638  
2642  
2646  
2650  
2654  
2658  
2662  
2666  
2670  
2674  
2678  
2682  
2686  
2690  
2694  
2698  
2702  
2706  
2710  
2714  
2718  
2722  
2726  
2730  
2734  
2738  
2742  
2746  
2750  
2754  
2758  
2762  
2766  
2770  
2774  
2778  
2782  
2786  
2790  
2794  
2798  
2802  
2806  
2810  
2814  
2818  
2822  
2826  
2830  
2834  
2838  
2842  
2846  
2850  
2854  
2858  
2862  
2866  
2870  
2874  
2878  
2882  
2886  
2890  
2894  
2898  
2902  
2906  
2910  
2914  
2918  
2922  
2926  
2930  
2934  
2938  
2942  
2946  
2950  
2954  
2958  
2962  
2966  
2970  
2974  
2978  
2982  
2986  
2990  
2994  
2998  
3002  
3006  
3010  
3014  
3018  
3022  
3026  
3030  
3034  
3038  
3042  
3046  
3050  
3054  
3058  
3062  
3066  
3070  
3074  
3078  
3082  
3086  
3090  
3094  
3098  
3102  
3106  
3110  
3114  
3118  
3122  
3126  
3130  
3134  
3138  
3142  
3146  
3150  
3154  
3158  
3162  
3166  
3170  
3174  
3178  
3182  
3186  
3190  
3194  
3198  
3202  
3206  
3210  
3214  
3218  
3222  
3226  
3230  
3234  
3238  
3242  
3246  
3250  
3254  
3258  
3262  
3266  
3270  
3274  
3278  
3282  
3286  
3290  
3294  
3298  
3302  
3306  
3310  
3314  
3318  
3322  
3326  
3330  
3334  
3338  
3342  
3346  
3350  
3354  
3358  
3362  
3366  
3370  
3374  
3378  
3382  
3386  
3390  
3394  
3398  
3402  
3406  
3410  
3414  
3418  
3422  
3426  
3430  
3434  
3438  
3442  
3446  
3450  
3454  
3458  
3462  
3466  
3470  
3474  
3478  
3482  
3486  
3490  
3494  
3498  
3502  
3506  
3510  
3514  
3518  
3522  
3526  
3530  
3534  
3538  
3542  
3546  
3550  
3554  
3558  
3562  
3566  
3570  
3574  
3578  
3582  
3586  
3590  
3594  
3598  
3602  
3606  
3610  
3614  
3618  
3622  
3626  
3630  
3634  
3638  
3642  
3646  
3650  
3654  
3658  
3662  
3666  
3670  
3674  
3678  
3682  
3686  
3690  
3694  
3698  
3702  
3706  
3710  
3714  
3718  
3722  
3726  
3730  
3734  
3738  
3742  
3746  
3750  
3754  
3758  
3762  
3766  
3770  
3774  
3778  
3782  
3786  
3790  
3794  
3798  
3802  
3806  
3810  
3814  
3818  
3822  
3826  
3830  
3834  
3838  
3842  
3846  
3850  
3854  
3858  
3862  
3866  
3870  
3874  
3878  
3882  
3886  
3890  
3894  
3898  
3902  
3906  
3910  
3914  
3918  
3922  
3926  
3930  
3934  
3938  
3942  
3946  
3950  
3954  
3958  
3962  
3966  
3970  
3974  
3978  
3982  
3986  
3990  
3994  
3998  
4002  
4006  
4010  
4014  
4018  
4022  
4026  
4030  
4034  
4038  
4042  
4046  
4050  
4054  
4058  
4062  
4066  
4070  
4074  
4078  
4082  
4086  
4090  
4094  
4098  
4102  
4106  
4110  
4114  
4118  
4122  
4126  
4130  
4134  
4138  
4142  
4146  
4150  
4154  
4158  
4162  
4166  
4170  
4174  
4178  
4182  
4186  
4190  
4194  
4198  
4202  
4206  
4210  
4214  
4218  
4222  
4226  
4230  
4234  
4238  
4242  
4246  
4250  
4254  
4258  
4262  
4266  
4270  
4274  
4278  
4282  
4286  
4290  
4294  
4298  
4302  
4306  
4310  
4314  
4318  
4322  
4326  
4330  
4334  
4338  
4342  
4346  
4350  
4354  
4358  
4362  
4366  
4370  
4374  
4378  
4382  
4386  
4390  
4394  
4398  
4402  
4406  
4410  
4414  
4418  
4422  
4426  
4430  
4434  
4438  
4442  
4446  
4450  
4454  
4458  
4462  
4466  
4470  
4474  
4478  
4482  
4486  
4490  
4494  
4498  
4502  
4506  
4510  
4514  
4518  
4522  
4526  
4530  
4534  
4538  
4542  
4546  
4550  
4554  
4558  
4562  
4566  
4570  
4574  
4578  
4582  
4586  
4590  
4594  
4598  
4602  
4606  
4610  
4614  
4618  
4622  
4626  
4630  
4634  
4638  
4642  
4646  
4650  
4654  
4658  
4662  
4666  
4670  
4674  
4678  
4682  
4686  
4690  
4694  
4698  
4702  
4706  
4710  
4714  
4718  
4722  
4726  
4730  
4734  
4738  
4742  
4746  
4750  
4754  
4758  
4762  
4766  
4770  
4774  
4778  
4782  
4786  
4790  
4794  
4798  
4802  
4806  
4810  
4814  
4818  
4822  
4826  
4830  
4834  
4838  
4842  
4846  
4850  
4854  
4858  
4862  
4866  
4870  
4874  
4878  
4882  
4886  
4890  
4894  
4898  
4902  
4906  
4910  
4914  
4918  
4922  
4926  
4930  
4934  
4938  
4942  
4946  
4950  
4954  
4958  
4962  
4966  
4970  
4974  
4978  
4982  
4986  
4990  
4994  
4998  
5002  
5006  
5010  
5014  
5018  
5022  
5026  
5030  
5034  
5038  
5042  
5046  
5050  
5054  
5058  
5062  
5066  
5070  
5074  
5078  
5082  
5086  
5090  
5094  
5098  
5102  
5106  
5110  
5114  
5118  
5122  
5126  
5130  
5134  
5138  
5142  
5146  
5150  
5154  
5158  
5162  
5166  
5170  
5174  
5178  
5182  
5186  
5190  
5194  
5198  
5202  
5206  
5210  
5214  
5218  
5222  
5226  
5230  
5234  
5238  
5242  
5246  
5250  
5254  
5258  
5262  
5266  
5270  
5274  
5278  
5282  
5286  
5290  
5294  
5298  
5302  
5306  
5310  
5314  
5318  
5322  
5326  
5330  
5334  
5338  
5342  
5346  
5350  
5354  
5358  
5362  
5366  
5370  
5374  
5378  
5382  
5386  
5390  
5394  
5398  
5402  
5406  
5410  
5414  
5418  
5422  
5426  
5430  
5434  
5438  
5442  
5446  
5450  
5454  
5458  
5462  
5466  
5470  
5474  
5478  
5482  
5486  
5490  
5494  
5498  
5502  
5506  
5510  
5514  
5518  
5522  
5526  
5530  
5534  
5538  
5542  
5546  
5550  
5554  
5558  
5562  
5566  
5570  
5574  
5578  
5582  
5586  
5590  
5594  
5598  
5602  
5606  
5610  
5614  
5618  
5622  
5626  
5630  
5634  
5638  
5642  
5646  
5650  
5654  
5658  
5662  
5666  
5670  
5674  
5678  
5682  
5686  
5690  
5694  
5698  
5702  
5706  
5710  
5714  
5718  
5722  
5726  
5730  
5734  
5738  
5742  
5746  
5750  
5754  
5758  
5762  
5766  
5770  
5774  
5778  
5782  
5786  
5790  
5794  
5798  
5802  
5806  
5810  
5814  
5818  
5822  
5826  
5830  
5834  
5838  
5842  
5846  
5850  
5854  
5858  
5862  
5866  
5870  
5874  
5878  
5882  
5886  
5890  
5894  
5898  
5902  
5906  
5910  
5914  
5918  
5922  
5926  
5930  
5934  
5938  
5942  
5946  
5950  
5954  
5958  
5962  
5966  
5970  
5974  
5978  
5982  
5986  
5990  
5994  
5998  
6002  
6006  
6010  
6014  
6018  
6022  
6026  
6030  
6034  
6038  
6042  
6046  
6050  
6054  
6058  
6062  
6066  
6070  
6074  
6078  
6082  
6086  
6090  
6094  
6098  
6102  
6106  
6110  
6114  
6118  
6122  
6126  
6130  
6134  
6138  
6142  
6146  
6150  
6154  
6158  
6162  
6166  
6170  
6174  
6178  
6182  
6186  
6190  
6194  
6198  
6202  
6206  
6210  
6214  
6218  
6222  
6226  
6230  
6234  
6238  
6242  
6246  
6250  
6254  
6258  
6262  
6266  
6270  
6274  
6278  
6282  
6286  
6290  
6294  
6298  
6302  
6306  
6310  
6314  
6318  
6322  
6326  
6330  
6334  
6338  
6342  
6346  
6350  
6354  
6358  
6362  
6366  
6370  
6374  
6378  
6382  
6386  
6390  
6394  
6398  
6402  
6406  
6410  
6414  
6418  
6422  
6426  
6430  
6434  
6438  
6442  
6446  
6450  
6454  
6458  
6462  
6466  
6470  
6474  
6478  
6482  
6486  
6490  
6494  
6498  
6502  
6506  
6510  
6514  
6518  
6522  
6526  
6530  
6534  
6538  
6542  
6546  
6550  
6554  
6558  
6562  
6566  
6570  
6574  
6578  
6582  
6586  
6590  
6594  
6598  
6602  
6606  
6610  
6614  
6618  
6622  
6626  
6630  
6634  
6638  
6642  
6646  
6650  
6654  
6658  
6662  
6666  
6670  
6674  
6678  
6682  
6686  
6690  
6694  
6698  
6702  
6706  
6710  
6714  
6718  
6722  
6726  
6730  
6734  
6738  
6742  
6746  
6750  
6754  
6758  
6762  
6766  
6770  
6774  
6778  
6782  
6786  
6790  
6794  
6798  
6802  
6806  
6810  
6814  
6818  
6822  
6826  
6830  
6834  
6838  
6842  
6846  
6850  
6854  
6858  
6862  
6866  
6870  
6874  
6878  
6882  
6886  
6890  
6894  
6898  
6902  
6906  
6910  
6914  
6918  
6922  
6926  
6930  
6934  
6938  
6942  
6946  
6950  
6954  
6958  
6962  
6966  
6970  
6974  
6978  
6982  
6986  
6990  
6994  
6998  
7002  
7006  
7010  
7014  
7018  
7022  
7026  
7030  
7034  
7038  
7042  
7046  
7050  
7054  
7058  
7062  
7066  
7070  
7074  
7078  
7082  
7086  
7090  
7094  
7098  
7102  
7106  
7110  
7114  
7118  
7122  
7126  
7130  
7134  
7138  
7142  
7146  
7150  
7154  
7158  
7162  
7166  
7170  
7174  
7178  
7182  
7186  
7190  
7194  
7198  
7202  
7206  
7210  
7214  
7218  
7222  
7226  
7230  
7234  
7238  
7242  
7246  
7250  
7254  
7258  
7262  
7266  
7270  
7274  
7278  
7282  
7286  
7290  
7294  
7298  
7302  
7306  
7310  
7314  
7318  
7322  
7326  
7330  
7334  
7338  
7342  
7346  
7350  
7354  
7358  
7362  
7366  
7370  
7374  
7378  
7382  
7386  
7390  
7394  
7398  
7402  
7406  
7410  
7414  
7418  
7422  
7426  
7430  
7434  
7438  
7442  
7446  
7450  
7454  
7458  
7462  
7466  
7470  
7474  
7478  
7482  
7486  
7490  
7494  
7498  
7502  
7506  
7510  
7514  
7518  
7522  
7526  
7530  
7534  
7538  
7542  
7546  
7550  
7554  
7558  
7562  
7566  
7570  
7574  
7578  
7582  
7586  
7590  
7594  
7598  
7602  
7606  
7610  
7614  
7618  
7622  
7626  
7630  
7634  
763